UNDERSTANDING PERFORMANCE MEASUREMENT IN THE CONTEXT OF HUMAN SERVICES DELIVERY PROGRAMS: IMPLICATIONS FOR YOUTH DEVELOPMENT PROGRAMS

Prepared for the
Family and Youth Services Bureau
by CSR, Incorporated

May 1999

TABLE OF CONTENTS

1. INTRODUCTION 1
2. UNDERSTANDING PERFORMANCE MEASUREMENT 2
2.1 What is performance measurement? 3
2.2 How can performance measurement information be interpreted and used? 5
3. DESIGNING AND IMPLEMENTING A PERFORMANCE MEASUREMENT SYSTEM: CHALLENGES ENCOUNTERED AND LESSONS LEARNED 7
3.1 Challenges to designing and implementing performance measurement in human services delivery programs 7
3.1.1 *Identifying appropriate and meaningful outcomes* 7
3.1.2 *Constructing valid and reliable outcome measures, data collection tools, and data tracking systems* 9
3.2 Factors facilitating development and implementation of performance measurement efforts 10
4. PERFORMANCE MEASUREMENT IN YOUTH DEVELOPMENT PROGRAMS 12
4.1 Identifying outcomes and measures for youth development programs 12
4.2 Developing a performance measurement system for programs funded by the Family and Youth Services Bureau 14

APPENDIX A: GUIDEBOOKS AND MANUALS FOR DEVELOPING PERFORMANCE MEASUREMENT SYSTEMS IN HUMAN SERVICES DELIVERY PROGRAMS 17

REFERENCES 20

UNDERSTANDING PERFORMANCE MEASUREMENT IN THE CONTEXT OF HUMAN SERVICES DELIVERY PROGRAMS: IMPLICATIONS FOR YOUTH DEVELOPMENT PROGRAMS

1. INTRODUCTION

In the arena of human services delivery programs, the term "performance measurement" refers to a procedure for documenting program or agency effectiveness. Traditionally, performance measurement in the human services arena has focused on documenting effectiveness in implementing intended interventions and services or meeting funders' requirements (Pratt, Katzev, Ozretich, Henderson, and McGuigan, 1998; Schorr, 1995). Recently, however, performance measurement has been used to also assess a program's effectiveness in achieving desired results for participants (Affholter, 1994; Hatry, 1997; Newcomer, 1997; Schorr, 1995). These efforts to assess participant results are not intended to replace assessments of program implementation. A meaningful performance measurement effort documents both program implementation and participant results (Pratt et al., 1998).

A performance measurement effort that focuses on participant results has clear benefits for human services delivery programs. Information about results can guide future program and policy planning and implementation (Pratt et al., 1998). In addition, information about results informs funders and the public about the consequences of their investments and minimizes investment in activities that are not productive of desired results (Schorr, 1995). Furthermore, targeting results broadens the definition of program "quality" to include not only how well the organization or program performs but also how responsive the organization is to its service population (Schorr 1995; Council on Quality and Leadership, 1998).

Documenting participant results may be particularly beneficial for youth development programs. These programs are designed to provide children and youth with the skills and supports they need to successfully navigate the journey from adolescence to adulthood (Dunkle, 1997; National Collaboration for Youth, 1997). The underlying assumption of youth development programs is that services for youth are more effective if they focus on youth's strengths and competencies rather than on their problems and deficiencies (Cahill and Pitts, 1997; CSR, Incorporated, 1996; Dunkle, 1997; Leffert, Benson, and Roehlkepartian, 1997). Although the general consensus of the youth services field is that a youth development approach is an effective framework for working with adolescents, there is little empirical evidence to support this assumption. In the present-day context of limited funding for youth development programs, empirical evidence that youth participating in these programs experience positive results may serve to promote their continued financial support (Dunkle, 1997; Leffert, Saito, Blyth, and Lroenke, 1996).

Despite the many benefits of results-focused performance measurement for human services delivery programs, it is not without risks. Even strong advocates recommend that developing a performance measurement system should be done carefully and cautiously (Schorr et al., 1995; Dunkle, 1997). Most human services delivery programs are designed to reach children and families with multiple problems and multiple service needs. Often, the changes that these children and families experience as a result of program participation are difficult to capture, particularly in the short-term (Schorr, 1995). Human services program advocates have expressed concern that results-focused performance measurement efforts, if not carefully designed and implemented, will engender further funding cutbacks for these programs because of the difficulty of demonstrating immediate "successes" (Dunkle, 1997; Schorr, 1995). To ensure that these efforts are designed and implemented correctly, performance measurement developers must have a clear understanding of what performance measurement is (and what it is not), and how information obtained from performance measurement can be appropriately interpreted and used (Bernstein, 1998; Hatry, 1997; Newcomer, 1996, 1997).

The purpose of this report is to present current information from the literature to support the appropriate design, implementation, and use of performance measurement in the human services delivery program arena. The information pertains to (1) an understanding of what performance measurement is, (2) the limitations of performance measurement information and how it can be interpreted and used correctly, and (3) the challenges and facilitating factors involved in developing and implementing performance measurement efforts in the human services program arena. The report also discusses the implications of this information for performance measurement efforts in the context of youth development programs and describes the efforts of the Family and Youth Services Bureau (FYSB) of the Department of Health and Human Services (HHS) to develop a performance measurement system for the youth development programs that it funds. An appendix to the report lists manuals and guidebooks that have been developed to assist various types of programs in establishing performance measurement systems.

2. UNDERSTANDING PERFORMANCE MEASUREMENT

In the previous section, performance measurement was briefly defined as a procedure for documenting program effectiveness. However, many other terms also refer to procedures for documenting program effectiveness, such as program evaluation, program monitoring, cost-benefit analysis, cost-effectiveness analysis, outcome evaluation, process evaluation, formative evaluation, summative evaluation, results-based accountability, outcome-based accountability, monitoring for accountability, and others. In fact, there is a striking lack of consistency in the field of program assessment regarding the ways that terms are defined and used (Freidman, 1997). Terms often appear in the literature without clear definitions. Some are used interchangeably, and others are defined in different ways across researchers. The multiplicity of terms, the use of various terms interchangeably, and the variation in definitions of similar terms contribute to the confusion regarding what performance measurement is and how performance measurement information can be used.

2.1 What is performance measurement?

Several experts in the field of program assessment believe that understanding what performance measurement is requires understanding how it is similar to and different from program evaluation (Bernstein, 1998; Fabry, 1998; Hatry, 1997; Newcomer, 1997; Pratt et al., 1998). One key similarity between the two assessment procedures is that both focus on a program's objectives. Program objectives are statements about what a program expects to accomplish in a given period of time (Broom and Jackson, 1997; Caudle, 1997). Objectives may refer to the resources to be allocated to a program (inputs); the products, activities, and services to be provided to participants (outputs); the number of participants to be reached (outputs); or the changes expected to occur among participants with respect to their knowledge, attitudes, behaviors, skills, or life circumstances (outcomes) (Pratt et al., 1998). The overarching goal of both performance measurement and program evaluation is to assess a program's success in achieving its objectives (Broom and Jackson, 1997; Duquette and Stowe, 1993; GAO, 1994; Hatry, van Houten, Plantz, and Greenway, 1996).

In assessing a program's objectives, however, performance measurement and program evaluation differ with regard to (1) the purpose of their assessments, (2) the kinds of questions asked, and (3) the research methodologies employed. Several program assessment experts suggest that the primary purpose of performance measurement is to monitor a program's progress over time (Fabry, 1998; Bernstein, 1998; Pratt et al., 1998; Wholey and Hatry, 1992). In fact, some authors use the terms "monitoring for accountability" or "program monitoring" interchangeably with "performance measurement" (Fabry, 1998; Hatry, 1992). As a monitoring tool, performance measurement examines the extent to which planned activities, outputs, and results are achieved and documents these achievements over time (Bernstein, 1998; Hatry, 1997; Hatry et al., 1996; Newcomer, 1997).

In comparison, the purpose of program evaluation is to provide a comprehensive understanding of a program and its affects on participants (Fabry, 1998; Bernstein, 1998; Rossi and Freeman, 1987). Program evaluation not only examines the extent to which objectives are achieved, but also the rationale for the development of objectives, the practices and procedures employed to attain objectives, and the interrelationships among objectives, specifically the contributions of resources and services to achieving expected participant outcomes (Hatry, 1997).

The different purposes of performance measurement and program evaluation are reflected in the kinds of questions each asks regarding program objectives. Performance measurement efforts typically ask the following questions (Pratt et al., 1998):

- What resources (inputs) are used to implement the program?
- What are the program activities and services (outputs)?
- What are the benefits for participants or for communities (outcomes)?
- What is the difference between actual inputs, outputs, and outcomes and those that were planned or required by funders?

A program evaluation asks these same questions, but also asks "how?" and "why?" (Newcomer, 1997). Why are programs not delivering the expected results? Why does implementation of the

same program vary across sites? How do specific program components contribute to outcomes achieved? Why are unintended negative and positive results occurring? Why are some members of the population more likely than others to achieve desired outcomes? How are outcomes related to resources (inputs) or interventions (outputs)?

Answering the "how?" and "why?" questions requires a different research methodology than answering the "what?" questions. To answer the "what?" questions, performance measurement efforts usually identify and quantify indicators or measures that demonstrate whether targeted levels of outputs or outcomes are attained (Bernstein, 1998; Hatry et al., 1996; Newcomer, 1997; Office of National Drug Control Policy, 1998). For example, a substance abuse intervention program for youth might plan to provide youth with a substance abuse education class (target output) to increase youth's knowledge of substance abuse information (target outcome). The program managers plan to provide 100 youth (output measure) with 50 class sessions (output measure) over the period of 1 year. The program expects that at the end of 1 year, at least 75 percent of the youth will achieve a score of 85 percent or better on a test of substance abuse knowledge (outcome measure). A performance measurement system applied to this program would document how many substance abuse education classes were actually provided during the year, the number of youth who actually participated in the classes, and the percentage of participating youth who achieved a score of 85 percent or better on a substance abuse knowledge test. By comparing actual measures to planned objectives, performance measurement can examine program success in achieving output and outcome objectives.

Although this methodology is sufficient to describe a program's accomplishments (or its failures), it is not sufficient to explain why or how they occurred (Affholter, 1994; Hatry et al., 1996; National Academy of Public Administration, 1994; Newcomer, 1997). A program evaluation methodology is necessary to provide this type of information. Program evaluations use social science research procedures such as experimental designs, random assignment of program participants to various experimental conditions, multiple assessments, and sophisticated statistical analyses (Bernstein, 1998; Fabry, 1998; Pratt et al., 1998). These research procedures are necessary to provide evidence that participant outcomes can be attributed to program operations or outputs and to describe the nature of other relationships between and among inputs, outputs, and outcomes.

For example, for the substance abuse intervention program described above, a program evaluation would not only provide more in-depth descriptive information about the substance abuse education classes and about the youth, but also would assess (1) whether participants' scores on a test of substance abuse knowledge given after the 52 class sessions differed significantly (as established by a statistical procedure) from their scores prior to the sessions, and (2) whether participants' post-session scores differed significantly from the scores of youth who did not receive the substance abuse education classes or who received a different type of substance abuse education class. This methodology permits inferences to be made regarding the relationship between the substance abuse classes and changes in participants' substance abuse knowledge.

2.2 How can performance measurement information be interpreted and used?

The disparate methodologies typically employed by performance measurement and program evaluation assessments have implications for interpreting and using resulting information. A concern frequently expressed by program assessment experts is that information obtained through performance measurement is being interpreted as representative of evaluation information, and inferences about causal relationships are being drawn without supporting evidence (Bernstein, 1998; Dunkle, 1997; Hatry, 1997; Newcomer, 1997). Performance measurement information can inform program managers and funders about program successes and problems. However, it cannot be used either to explain the reasons for successes or problems or to provide indications of how problems might be resolved (Bernstein, 1998; Newcomer, 1997). In contrast, program evaluation information, if the correct methodology is employed, can be used to make inferences regarding the relationships between outputs and outcomes and consequently can explain successes and point to potential solutions for identified problems (Bernstein, 1998; Fabry, 1998; Newcomer, 1997; Pratt et al., 1998).

For most program assessment experts, the primary function of performance measurement information is to enhance management decisionmaking (Bernstein, 1998; Joyce, 1997; Wholey, 1997). As a management tool, performance measurement information can be used to make decisions regarding allocations of resources (Joyce, 1997; Wholey, 1997), redirect program activities (Wholey, 1997), and support policy decisionmaking (Wholey, 1997). A key decision that might be made on the basis of performance measurement information is that a program evaluation is needed to further explain problems and to understand ways that problems might be resolved. To use the substance abuse intervention program for youth as an example, a performance measurement assessment might provide program managers with the following information:

> In program year 1, 52 substance abuse education classes were provided (1 per week) to 100 youth. Of those 100 youth, 67 percent scored at 85 percent or better on a substance abuse knowledge test taken at the end of the year. The program expected that at least 75 percent of participants would score 85 percent or better on the test.

This information tells program managers that the program has been successful in attaining planned outputs (i.e., number of classes and number of youth reached), but has not been successful with respect to attaining the expected level of performance (i.e., 75 percent of participants achieving a score of 85 or better on the test) that would demonstrate the program's success in reaching the desired outcome (increased knowledge of substance abuse-related information). Since this may have been a key outcome for the program, managers or funders may decide that the program is a failure.

In fact, performance measurement information alone is inadequate to support the conclusion of program failure (Affholter, 1994; Bernstein, 1998). Such a conclusion can only be drawn from the results of a program evaluation that examines the relationship between the substance abuse education classes and the test scores and various aspects of the education component (curriculum, educational approach, characteristics of instructors, etc.) and participants (demographic characteristics, literacy levels, prior knowledge, etc.).

For example, a program evaluation might discover that even though less than 75 percent of the youth achieved a score of 85 or better on the test of substance abuse knowledge (expected outcome), the youth's test scores after the class sessions were significantly higher than their scores prior to class participation. In addition, the youth participating in the class scored significantly higher on the test than a control group composed of youth who did not participate in the class. This would suggest that the education class did have a positive affect on youth's substance abuse knowledge.

The next question for a program evaluation, then, is why the expected performance measure (75 percent of youth scoring 85 percent or better on the test of substance abuse knowledge) was not achieved. Program evaluation can assess whether the problem can be attributed to characteristics of the educational process itself, characteristics of the youth, characteristics of the measurement/ assessment tool, or to some interactive combination of these factors. A program evaluation might discover, for example, that youth who have higher literacy levels performed better than "low-literacy" youth on the test both prior to and after participating in the classes. This implies that the literacy level of the items on the knowledge test may be too advanced for all youth in the target population. It may be possible to resolve this problem by revising the test items so that they are more easily understood by all target population members. This could result in more of the youth meeting the performance expectation of an 85 percent test score the next time that the program's performance is assessed.

This example suggests that although performance measurement can document inputs, outputs, and outcomes, program evaluation is necessary to understand how inputs and outputs are related to outcomes (and to one another), and identify ways that performance problems might be resolved (Affholter, 1994; Bernstein, 1998; Newcomer, 1997). The example also suggests that ultimate decisions about a program's value with respect to its affects on participants cannot be made solely on the basis of performance measurement information (Bernstein, 1998; Newcomer, 1997). Instead, a comprehensive picture of program success requires both performance measurement and program evaluation information (Bernstein, 1998; GAO, 1997a; Pratt et al., 1998). The need for both types of information was recognized in the Federal Government's Government Performance and Results Act of 1993 (GAO, 1997b). This Act requires Federal agencies to provide a description of previous program evaluation information and a schedule for future program evaluations in their annual performance plans, and to include a summary of program evaluation findings as well as performance measurement results in their annual performance reports (Bernstein, 1998).

3. DESIGNING AND IMPLEMENTING A PERFORMANCE MEASUREMENT SYSTEM: CHALLENGES ENCOUNTERED AND LESSONS LEARNED

When performance measurement in human services delivery programs focused primarily on program inputs and outputs, the task of designing and implementing a performance measurement effort was relatively straightforward. The incorporation of outcomes as part of performance

measurement has greatly increased the complexity of this task. Several reviews of outcome-based performance measurement of human services delivery programs have identified key challenges and facilitating factors to designing and implementing these efforts (Epstein and Olsen, 1996; GAO, 1997a; Hatry, 1997; American Humane Association, 1994; Plantz, Greenway, and Hendricks, 1997). These are discussed in the following sections.

3.1 Challenges to designing and implementing performance measurement in human services delivery programs

The most significant challenges to designing and implementing an outcome-based performance measurement effort for human services delivery programs are (1) identifying appropriate and meaningful outcomes, (2) constructing valid and reliable measures for assessing outcomes; and (3) developing effective data collection tools and tracking systems (Hatry et al., 1996; American Humane Association, 1994; Plantz et al., 1997).

3.1.1 *Identifying appropriate and meaningful outcomes*

An outcomes-based performance measurement system requires that the outcomes selected to assess performance must be meaningful with respect to program operations. That is, the outcomes must (1) be consistent with agency or program missions and goals; (2) meet the needs of stakeholders, (3) reflect agency responsibility, and (4) pertain to a desired result for program participants or for the program (Pecora, 1998).

Identifying outcomes that meet all of these criteria is a challenging task for human services delivery programs. In fact, there is a lack of consensus in the field as to the types of outcomes that should be the focus of a human services program's performance measurement system. Some advocates are adamant that meaningful assessments of program effectiveness require a focus on outcomes that address changes in the program participants themselves (The Council, 1998; Curtis, 1996; Dunkle, 1997). Did the youth attain high school diplomas? Did the parents acquire positive parenting skills? Are children performing at expected levels in school? Are youth avoiding high-risk behaviors? Other advocates suggest that programmatic outcomes, which describe how well a program is doing with respect to attaining its objectives for participants, are also meaningful assessments of program effectiveness (Curtis, 1996). These outcomes may pertain to events such as increases in the number of youth served by runaway and homeless youth programs or by child welfare programs who are reunified with their parents, decreases in the amount of time children spend in foster care, increases in the number of street youth seeking services from community agencies, and the like.

There also is a lack of consensus in the field as to whether outcomes selected for a performance measurement effort should be "immediate," "intermediate," or "longer-term" (Hatry et al., 1996; Newcomer, 1997; Plantz et al., 1997). Immediate (or initial) outcomes are those that are expected to occur as a direct result of program participation, such as an increase in youth's substance abuse knowledge as a result of participation in a substance abuse education class. These outcomes may not be especially meaningful in terms of the quality of participants' lives, but they are assumed to

represent an individual's progress toward a more meaningful change. Intermediate (or interim) outcomes are expected to occur as a result of the changes reflected by immediate outcomes. For example, the knowledge gains experienced by youth as a result of the substance abuse intervention program (immediate outcomes) are expected to produce changes in youth's substance use behaviors, specifically a decrease or cessation of use (intermediate outcomes). Longer-term outcomes generally reflect the more substantive life changes that are expected to occur once individuals have achieved the intermediate outcomes. An example of this is youth demonstrating long-term academic success as a result of their cessation of substance use. Some program assessment experts suggest that it is beyond the scope of a performance measurement system to document long-term outcomes because of the difficulty and expense associated with tracking participants for long periods of time. However, others suggest that documenting meaningful change requires including outcomes in each of the three categories (immediate, intermediate, and longer-term) (Kirk, 1994).

One potential guideline for determining the level of outcome to incorporate into a performance measurement system is the amount of empirical evidence supporting the hypothesized relationships between immediate and intermediate outcomes and between intermediate outcomes and longer-term outcomes. For example, increases in substance abuse knowledge may be a meaningful outcome for a performance measurement system if there is empirical evidence to support the assumption that gains in substance abuse knowledge are related to decreases in substance use. If there is no empirical evidence to support this linkage, then attaining the outcome of a change in substance abuse knowledge may be meaningless in terms of improving the status of youth participating in the program. When there is inadequate empirical support linking immediate outcomes to other levels of outcomes, the focus of the performance measurement effort should be on intermediate or longer term outcomes.

Selecting outcomes for human services delivery programs also is problematic because the most meaningful outcomes may be those that are influenced by factors beyond the control of the program or agency. It would be difficult for a substance abuse education class, for example, to eventually result in youth decreasing or ceasing their substance use over time unless there are other resources in the community, such as Boys and Girls Clubs or substance abuse treatment programs, that provide continued support for the initial behavior changes.

Another reason why selecting meaningful outcomes for a performance measurement effort is a challenging task is that many outcomes that are meaningful may not be measurable. In fact, a general concern expressed in the literature is that for many performance measurement efforts, the outcomes selected will be those that are easiest to measure rather than those that are the most meaningful (Dunkle, 1997; Newcomer, 1997). This issue is particularly relevant for human services delivery programs that do not target specific problem behaviors but instead provide multiple services and activities in an effort to enhance the general life conditions of the children, youth, or families they serve. The services might include youth development programs, family support programs, home nursing programs, and the like. Often the kinds of changes that occur as a result of these latter types of programs are incremental and difficult to measure, and truly meaningful outcomes may not be observable for a long time. The challenges inherent in measuring outcomes and accessing the necessary data are discussed below.

3.1.2 *Constructing valid and reliable outcome measures, data collection tools, and data tracking systems*

Outcome measures or indicators (these terms are sometimes used interchangeably) are specific items of data that are tracked to assess the extent to which a program is achieving an outcome (Hatry et al., 1996). To use the substance abuse intervention program as an example, the indicator or outcome measure is the youth's score on the test of substance abuse knowledge. This score is used to assess the extent to which the program is achieving its desired immediate outcome of increasing youth's knowledge of substance abuse information. A measure or indicator that would assess attainment of the intermediate outcome of decreasing youth's substance use behaviors might be youth's self-report of current use compared to self-report of use prior to program participation.

Constructing outcome measures for which data are either readily available or can be collected without extensive cost or additional burden to program staff is a difficult process. Theoretically, outcome measures must be reliable, valid, relevant to the local program context, feasible to implement, and applicable across cultures (Pecora, 1998). One researcher (Friedman, 1997) suggests that an outcome measure should have "communication power" (it is compelling and clear), "proxy power" (there is a strong relationship between the measure and the outcome that has been established either by research or common sense), and "data power" (the indicator/measure is assessed by high-quality, regular, and frequently reported data). In reality, outcome measures in the arena of human services delivery programs rarely meet these criteria.

The criterion that is most difficult for human services delivery programs to meet concerns the availability of high-quality, regular, and frequently reported data. One reason for this is that data collection or assessment tools often are not sufficiently sophisticated to meet the outcome measurement needs of human services delivery programs. In many of these programs, target populations may exhibit very small (but important) changes over the course of program participation that are difficult to measure in a reliable manner (Malone and Markley, 1998). Tools for collecting data on these changes often are lacking, difficult to interpret, or too primitive for the task (Dunkle, 1997; Malone and Markley, 1998). Consequently, the most meaningful data, particularly for individual programs, often are not collected (Dunkle, 1997).

Access to high-quality, regular, and frequently reported data also is challenging for many human services delivery programs because it often requires a computerized information system with the capacity to collect, store, analyze, track, and report data pertaining to specified outcomes (Caudle, 1997; Malone and Markley, 1998). Establishing this type of system can be a complex, time consuming, and expensive operation. Many human services delivery programs lack the technical expertise and/or the necessary financial resources to establish a data collection and management information system.

Despite the challenges encountered, efforts to establish performance measurement systems within the context of human services delivery programs have met with some success. The factors that were found to facilitate successful implementation are presented below.

3.2 Factors facilitating development and implementation of performance measurement efforts

Recent efforts to develop and implement performance measurement systems in the context of human services delivery programs have been undertaken by a variety of organizational entities ranging from Federal agencies within HHS, to State child welfare systems, to private nonprofit community-based services agencies. Despite the variation in organizational types, reviews of these efforts reveal a high degree of similarity with respect to the kinds of factors that facilitate successful implementation (Epstein and Olsen, 1996; GAO, 1995; Malone and Markley, 1998; Wholey and Newcomer, 1997). The following are the key findings of these reviews with respect to facilitating factors:

- The development of a performance measurement system is facilitated by a clear vision, statement of purpose, and a set of goals that are maintained throughout the implementation process (Wholey, 1997). An effective statement of purpose and goals addresses (1) why the system is necessary and how it fits into the program's decision making process; (2) how performance information will be used; and (3) the ways that performance measurement can support program staff and activities (Broom and Jackson, 1997; Epstein and Olsen, 1996; GAO, 1995; Wholey, 1997).

- Although it is important that all levels of program staff be committed to the performance assessment process (Epstein and Olsen, 1996; GAO, 1995), leadership at the senior levels is the critical element that can make or break performance measurement efforts (Malone and Markley, 1998; Wholey and Newcomer, 1997). Commitment to the performance measurement process is facilitated when everyone has the opportunity to have input into the design and development of the system and when technical assistance is provided in developing outcomes and measures (Malone and Markley; 1998).

- The effectiveness of a performance measurement system is enhanced when sufficient time is allowed for the development and implementation process. Because performance measurement is a system of documenting program effectiveness over time, designing and testing the system can require several years. Leadership and staff must be willing to make a long-term commitment (Epstein and Olsen, 1996; GAO, 1995) and accept that measures will be imperfect at the beginning of the process, then gradually improve as they are pilot tested and refined (Epstein and Olsen, 1996).

- Implementation of a performance measurement system is facilitated when the system itself is designed as a long-term system—that is, the system can withstand apparent short-term program failures, is readily adaptable to changing circumstances, and permits changes in focus as needed during the course of implementation (Broom and Jackson, 1997; Epstein and Olsen, 1996).

- Training of program managers, staff, and other stakeholders in the performance measurement system is critical for successful implementation (Broom and Jackson, 1997; National Academy

of Public Administration, 1994). Training is needed to familiarize stakeholders with the basic concepts and terminology of performance measurement so that there is a shared understanding of the system. The most effective training is hands-on and practical, particularly with respect to collecting the data needed to measure outcomes and/or building and maintaining computerized information systems to store, track, analyze, and report performance measurement information.

- Implementation is facilitated when the measures to be assessed are limited to a vital few that provide the most needed information with respect to documenting outputs and outcomes and making policy and program decisions (Wholey, 1997). Too many measures can confuse and overwhelm users or make a performance measurement system unmanageable (GAO, 1995; Malone and Markley, 1998).

- Performance measurement efforts are fostered when there is regular reporting of performance data and when reports are circulated among the stakeholders who are involved in and are affected by the performance measurement system (Epstein and Olsen, 1966).

- The development of a performance measurement system is enhanced when program evaluation information is used to assist in the development of measures for assessing whether outcomes have been attained. Program evaluation information can serve as a guide for establishing the levels of performance that are both desirable and realistic. (GAO, 1995).

- When designers of a performance measurement system begin their efforts by learning about what others in similar circumstances have done, the entire process is facilitated and the potential for costly and time-consuming mistakes is reduced (Plantz et al., 1997).

4. PERFORMANCE MEASUREMENT IN YOUTH DEVELOPMENT PROGRAMS

An understanding of performance measurement and the challenges and facilitators to implementation provides a useful structure for understanding and implementing performance measurement efforts for youth development programs. The challenges associated with selection of outcomes, development of measures, and access to data and data collection tools are particularly pertinent to youth development programs (Dunkle, 1997).

4.1 Identifying outcomes and measures for youth development programs

As noted previously, youth development programs approach youth in terms of their strengths and resources rather than their problems and deficits (Dunkle 1997). These programs focus on fostering positive developmental experiences as the primary route to positive outcomes (Cahill and Pitts, 1997; Perkins, Haas, and Keith, 1997). Because of this approach, youth development programs do not frame expected objectives for youth in terms of avoidance or reduction of negative behaviors, such as abstaining from substance abuse, not getting pregnant as a teenager, or not engaging in other high-risk behaviors. Instead, these programs tend to describe their objectives

for youth in terms of positive growth, such as improvements in school performance, improved relationships with parents and/or peers, increased knowledge and skills pertaining to employment or independent living, increased self-sufficiency, and the like.

A key challenge for developing an outcomes-based performance measurement system for youth development programs, then, is identifying outcomes that assess change from the perspective of positive growth and development Many youth development programs have approached this challenge by first determining the kinds of characteristics of youth that have been reported in the research literature to be associated with positive growth and development during adolescence and then translating these characteristics into meaningful outcomes for assessing program success.

The several research reviews designed to identify the youth characteristics associated with positive development during adolescence have produced highly consistent findings. For example, HHS developed a delinquency prevention program in 1970 based on findings of a research review indicating that the four ingredients youth need to develop in a positive way are a sense of competence, a sense of usefulness, a sense of belonging, and a sense of personal power (ACF, 1966). More recently, a FYSB-sponsored review of the literature on adolescent development (CSR, 1996) reported that the four ingredients youth need to develop in a positive way are a sense of competency, a sense of connectedness to others and to society, a belief that they can exert some control over their fate in life, and a stable sense of identity. Although the terms used in the two reports are somewhat different (e.g., sense of belonging rather than a sense of connectedness, and sense of power rather than sense of control), the findings are nearly identical despite the fact that the literature reviews were conducted almost 30 years apart. Similar findings were reported in research efforts conducted by Public/Private Ventures (Gambone, 1997; Gambone and Arbreton, 1997) and the Search Institute (Benson, 1993; Leffert et al., 1966).

Several organizations have translated these “ingredients” into meaningful outcomes and outcome measures for youth development programs. The Center for Youth Development and Policy Research of the Academy for Educational Development (1996), for example, recommends that positive development could be assessed by measures that determining whether a youth is:

- Able to identify three recent accomplishments (competency);
- A member of an organization or club, or involved in community volunteer activities (belonging or connectedness to society);
- Able to identify a book he or she has read for pleasure in the past 2 months;
- Consistently involved in school or a training program (competency, power); and/or
- Consistently connected to at least one caring adult (belonging, connectedness).

The Search Institute developed a model that incorporates 40 developmental assets, many of which are linked to the factors of competency, connectedness/belonging, and power/control. A recent Search Institute report (Scales and Leffert, 1999) reviewed over 800 studies that measured the effects of developmental assets in youth and reported, among other findings, that supportive and caring relationships (connectedness) are more fundamental than program interventions in building assets.

Youth development programs operated under the auspices of Girls Incorporated (Frederick and Nicholson, 1991) are encouraged to assess the following expected outcomes as part of their performance measurement assessments:

- Increases in knowledge about key areas of their lives both now and in the future (control and competency);
- Improvements in school or academic achievement (competency);
- Positive outlooks and values (control);
- Improvements in relationships with peers and parents (connectedness); and
- Improvements in decisionmaking and problemsolving (competency).

The Girl Scouts of the U.S.A. has focused on the following expected outcomes in its youth development programs (Hwalek and Minnick, 1997), most of which are related to competency, connectedness/belonging, and power/control:

- Self reliance (independent living skills);
- Self competence;
- Social skills;
- Respect for others;
- Feelings of belonging;
- Values clarification/decisionmaking;
- Helpfulness/concern for community; and
- Teamwork and leadership.

Although there is consistency with regard to the types of outcomes that may provide meaningful assessments of youth development programs, there is less consistency with regard to how these outcomes can be reliably and validly measured. Measuring outcomes such as respect for others, values clarification, and improvements in relationships with parents and peers requires sophisticated observation or assessment tools. These tools must not only be capable of assessing the outcome measure but also must be sensitive enough to capture changes in individuals with respect to the measure, changes that in some instances may be small and difficult to capture. Although some national programs have developed their own assessment tools, the usefulness and validity of these tools has not yet been well established.

4.2 Developing a performance measurement system for programs funded by the Family and Youth Services Bureau

In addition to private organizations such as the Search Institute, Girl Scouts, and Girls Incorporated, the Federal Government has been involved in efforts to develop a performance measurement system for the youth development programs that it funds through the activities of FYSB. FYSB's mission is to provide national leadership on issues concerning the welfare of youth and to assist individuals and organizations in providing effective, comprehensive services for youth in at-risk situations and their families. FYSB encourages communities to support young people through a youth development approach.

Under the Runaway and Homeless Youth Act, FYSB provides funding to over 350 grantees operating one or more of the following three types of youth development programs:

C *Basic Center Program:* This program supports youth shelters that provide emergency shelter, food, clothing, outreach services, and crisis intervention for runaway and homeless youth. The shelters also offer services to help reunite youth with their families whenever possible.

C *Street Outreach Program*: This program funds local youth service providers to conduct street-based education and outreach and offer emergency shelter and related services to young people with the goal of helping them leave the streets.

C *Transitional Living Program*: This program assists older, homeless youth to develop skills and resources necessary for independent living. Housing and services are provided for up to 18 months to youth between the ages of 16 to 21 who are unable to return to their homes.

During the past year, FYSB, drawing on the lessons of other performance measurement efforts, has developed a performance measurement system incorporating a set of outcomes and outcome measures that can be used to assess the performance of its youth development programs. The primary goal of FYSB's performance measurement efforts is to identify areas of success and areas where improvements may be needed. The performance measurement system is based on the following principles and components:

- **The system reflects the goals and methodology of performance measurement as compared to program evaluation.** The FYSB outcomes-based performance measurement system is designed to track the extent to which programs achieve specified results for participants. Information from the system will be used to enhance management decisionmaking regarding program operations and identify areas where technical assistance or further assessment may be needed. FYSB recognizes that a performance measurement system is not a program evaluation, and that performance measurement information cannot be used to explain why results were or were not attained, or to establish causal relationships between program activities/interventions and participant outcomes. Therefore, it is acknowledged that information from performance measurement alone is not sufficient to make decisions about a program's ultimate value.

- **The system is designed to foster continuous improvement of individual programs.** FYSB *will not* establish a universal standard of success for each measure and compare an individual program's performance to that standard. Because of the extensive variation in types of populations served, types of grantee agencies, and duration and intensity of services across FYSB grantee programs, the development of a universal performance standard is not feasible. Instead, FYSB's performance measurement system is designed to promote continuous improvement by tracking an individual program's change in performance over time. The primary performance system question addressed will be: From one measurement time (e.g., a year) to the next, did a program's performance on a particular measure improve, show no change, or decline?

- **The system incorporates a limited number of key outcomes that are significant to both FYSB and the grantees.** During the consultation process, stakeholders identified an extensive number of potential outcomes. However, one of the "lessons learned" reported in the performance measurement literature is that too many outcomes can confuse and overwhelm users and make the system unmanageable. Therefore, in selecting outcomes from those generated by stakeholders, FYSB focused on those that provide the most important information for programs in terms of participant results.

- **The selected outcomes are consistent with FYSB's primary goal of enhancing youth well-being.** FYSB's performance plan for the Government Performance and Results Act lists one of its major goals as enhancing youth well-being. This goal incorporates the following two components: (1) the safety of youth at home or in appropriate alternative settings, and (2) the progress of youth in their development into independent, contributing members of society. With respect to the latter component, research reviews have been highly consistent in their reports regarding characteristics that are associated with positive development during adolescence. The three consistently identified characteristics that were viewed as particularly relevant to FYSB-funded programs are: (1) a sense of connectedness to others and to society (also described as a sense of belonging); (2) a sense of competency and industry; and (3) a sense of control over one's fate in life (also described as a sense of power). These characteristics, in addition to the component of youth safety, provided the conceptual domains for outcome selection.

- **Outcome measures were selected on the basis of (1) the meaningfulness of the measure with respect to the selected outcomes, and (2) the feasibility of collecting valid and reliable data pertaining to the measure.** FYSB established that to be included in the system, a measure must not only be meaningfully linked to a desired outcome, but also must be assessable through an item or a set of items that can be added to an existing data collection system. In addition, the measures should reflect immediate or intermediate expected outcomes rather than longer-term outcomes, because, as many stakeholders indicated, tracking program participants over extended periods of time is beyond the data collection capabilities of many programs. However, some longer-term outcome measures are included as options for those programs that already have longer-term tracking systems or are planning to implement these systems.

The following outcomes were selected as the key changes that are expected to occur:

- Youth will be in a stable and safe living situation;
- Youth's perceptions of their relationships with positive adults and peers will improve;
- Youth's interest in community involvement will increase;
- Youth's perceptions of their personal abilities will improve; and
- Youth's perceptions that they can affect their futures will improve.

FYSB's performance measurement system has not yet been pilot tested or implemented. However, the efforts of FYSB and other national and local organizations to develop performance measurement systems for youth development programs represents a significant step toward helping these programs make the case for their continued existence and financial support (Leffert et al., 1996). In addition, the fact that many of these efforts are based on research findings regarding the characteristics of youth associated with positive developmental pathways means that directions taken in developing the performance measurement system are supported by empirical evidence rather than based on assumptions or beliefs. However, despite the successes of these efforts, it is important that FYSB, other national organizations, and program funders keep in mind that performance measurement information only provides descriptive information regarding program results. Information about why and how these results occurred can come only from program evaluations. Consequently, decisions regarding a program's ultimate worth cannot be made on the basis of performance measurement information alone.

APPENDIX A:
GUIDEBOOKS AND MANUALS FOR DEVELOPING PERFORMANCE MEASUREMENT SYSTEMS IN HUMAN SERVICES DELIVERY PROGRAMS

Measuring Program Outcomes: A Practical Approach. (Hatry, van Houten, Plantz, and Greenway, 1996. Alexandria, VA: United Way of America.) This United Way-sponsored manual provides detailed guidance to public and nonprofit organizations regarding the development and implementation of an outcome-based performance measurement system. The information in the manual was based on more than a decade's experience in developing and studying outcome-based program performance measurement systems and on input from a variety of stakeholders across many different types of human services programs. The manual offers a step-by-step process for identifying outcomes, indicators, and data collection methods and for reporting and using data. Checklists and examples of useful worksheets are available to assist in developing an outcome measurement system. However, program-specific outcome indicators or data collection methods are not provided. The manual also incorporates quotes and vignettes from more than 30 agencies that have implemented programs using results of outcome-based performance measurement.

Chapter Management Workbook-Performance Measurement: The Development of Outcome Objectives. (The American Red Cross, 1996–1997, Washington, DC) The American Red Cross developed workbooks designed to help local chapters develop outcome objectives and appropriate indicators. Five workbooks are available, one for each line of service provided by the Red Cross (Armed Forces Emergency Services, Blood Services, Disaster Services, Health and Safety Services, and International Services). The workbooks present step-by-step processes for defining program goals and objectives, selecting data collection methods, selecting indicators, and identifying outcome targets.

Training Manual for Human Service Organizations Interested in Developing Outcome-Focused Program Performance Measurement Systems. (Robertson, 1996. Allentown, PA: Valley Youth House.) Robertson's guide is designed to help programs transition from the documentation of *process* (counting clients and services) to the documentation of *outcomes* (counting measurable changes in behaviors) for demonstrating success and accountability. This manual was designed to provide human service organizations and staff with an introduction to the concepts, strategies, and procedures of outcome-based program performance measurement. It includes sections on benefits of outcome-based systems for measuring the success of a human services program; concepts and procedures with regard to outcomes and outcome measures; extensive guidelines for implementing an outcome measurement system (accountability measures); and pitfalls in designing and implementing an outcome measurement system. The manual also includes many examples of measures that human service programs can use or adapt in the process of designing their own outcome measurement system. Categories of measures include individual measures of pretest and posttest levels of functioning, condition, or status; goal attainment scales; client satisfaction surveys; and the Valley Youth House Staff Service Report. This publication

includes some outcome measures and goal attainment sections. It provides a positive, user-friendly approach to presenting basic concepts and strategies for moving from providing anecdotal, unsubstantiated success stories to demonstrating success with data.

Assess for Success: Needs Assessment and Evaluation Guide. (Frederick and Nicholson, 1991. Indianapolis, IN: Girls Incorporated.) This guide offers instructions, recommendations, and sample forms for designing, implementing, and evaluating Girls Incorporated programs. The section on program evaluation discusses three levels of evaluation: process evaluation, outcome or impact evaluation, and scientific outcome or impact evaluation. It presents three common areas of outcome measurement (knowledge, attitudes, and behaviors) and ways to measure them. The guide also discusses data collection for evaluating programs. It reviews the use of attendance records, program descriptions, staff logs, expert observers, surveys, certifications and informal contracts, and unobtrusive techniques, such as looking for physical evidence of changes or significant choices, developing role-playing situations in which girls must demonstrate knowledge or other outcomes, and looking for changes in the physical environment. The guide lists several criteria for selecting outcome measures, such as those that demonstrate program success and that are the least disruptive, those that are practical and affordable to measure, and those that have the capacity to engage participants. Outcome measures also should be straightforward, concise, and consistent with each other.

Girl Scout Outcomes Study Procedures Manual. (Hwalek and Minnick, 1997. New York: Girl Scouts of the U.S.A.) This is a detailed procedures manual available to Girl Scout councils wanting to conduct a local Girl Scout outcomes study. It provides reasons why outcome measurement has become a priority among programs for children. Issues are discussed regarding the use and potential misuse of outcome measures with children. The research design for conducting a Girl Scouts outcomes study also is presented. Procedures are recommended for determining sample size and for selecting study participants. Responsibilities of staff are discussed, and all necessary letters, forms, instructions, and other items that a council would need to conduct the local study are included in both hard copy and computer diskette formats.

Scouting's Positive Impact on the Community: A Resource for Measuring Program Outcomes. (Boy Scouts of America, 1996. Irving, TX: Boy Scouts of America.) This booklet was designed to support local councils in developing their United Way presentations for outcome measures funding for impact and evaluation. It defines outcome measures as "problem-related, attainable, and measurable statements of a program's intended effects on the knowledge, skills, behavior, or condition of those it is designed to help." The booklet contains (1) United Way community objectives; (2) BSA programs and activities; (3) values that each boy/youth receives in the Scouting program; and (4) impact and outcome results of a boy's/young adult's positive participation in Scouting/Learning for Life for the betterment of the community.

The Handbook of Positive Youth Outcomes. (Youth Development Institute, 1998. New York: Fund for the City of New York.) This handbook was created by the Networks for Youth

Development to enable youth agencies "to share insights and tap into each other's expertise." Presented in a pocket-sized format, this handbook outlines the necessary outcomes for youth in effective programming as determined by the contributing agencies. The outcomes are divided among the following eight broad areas of competence: originality (creative competency), understanding ourselves and others (personal competency), thinking and reasoning (cognitive competency), civic competency, our bodies (physical health competency), mental health competency, employability competency, and social competency. Presented within these eight areas of competence are more detailed lists of outcomes. The handbook can also be reversed to view a regrouping of the outcomes into skill areas necessary for work, educational, social, and personal lives. The intent of the handbook is to help programs define, structure and evaluate their work with young people within the framework of positive youth development.

Personal Outcome Measures for Children and Youth. (The Council on Quality and Leadership in Supports for People with Disabilities, 1998, Towson, MD: The Council.) The Council's goal is to ensure that children with disabilities have the opportunity to lead full and abundant lives. This manual serves as a guide in using 25 personal outcomes to enhance services provided to children and youth with disabilities. It defines the outcomes, identifies core practices that support the outcomes, and supplies a process for discovering whether the outcomes are present for each youth.

REFERENCES

Affholter, D.P. 1994. "Outcome Monitoring." In J. S. Wholey, H.P. Hatry, and K.E. Newcomer (eds.). *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass.

American Humane Association. 1994. *A Framework for Outcome Measures in Child Welfare Services*. Paper prepared for the Second Annual Roundtable on Outcome Measures in Child Welfare Services, San Antonio, TX, April 7–9. Washington, DC: American Human Association.

Benson, P. L. 1993. *The Troubled Journey: A Portrait of 6th–12th Grade Youth*. Minneapolis: Search Institute.

Bernstein, D.J. 1998. "The Theoretical and Potential Relationship of Performance Measurement and Program Evaluation." Presentation at the Annual Meeting of the American Evaluation Association.

Broom, C.A., and Jackson, M. 1997. "Performance Measurement Training That Works." In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:79–90.

Cahill, M., and Pitts, L. 1997. "Strengthening Youth Employment Prospects Through Youth Development." Unpublished paper. New York: Youth Development Institute, Fund for the City of New York.

Caudle, S.L. 1997. "Performance Results: The Information Technology Factor." In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:63–78.

Center for the Study of Social Policy. 1995. *Building New Futures for At-Risk Youth: Findings from a Five-Year, Multi-Site Evaluation.* Washington, DC: Center for the Study of Social Policy.

Center for Youth Development and Policy Research, Academy for Educational Development. 1996. *Youth, Program, and Community Outcomes for Youth Development*. Washington, DC: Academy for Educational Development.

The Council on Quality and Leadership in Supports for People with Disabilities. 1998. *Personal Outcome Measures for Children and Youth.* Field Review Edition. Towson, MD: The Council.

CSR, Incorporated. 1996. *Understanding Youth Development: Promoting Positive Pathways of Growth*. Prepared for FYSB/ACYF/ACF/HHS. Washington, DC: U.S. Government Printing Office.

Curtis, P.A. 1996. "Outcomes in Child Welfare." Paper presented to the Hudson River Consortium, West Point, NY, April 19.

Dunkle, M.C. 1997. *Steer, Row or Abandon Ship? Rethinking the Federal Role for Children, Youth and Families*. Washington, DC: The Institute for Educational Leadership.

Duquette, D.J., and Stowe, A.M. 1993. "A Performance Measurement Model for the Office of Inspector General." *Government Accountants Journal* Summer:27–50.

Epstein, J., and Olsen, R.T. 1996. "Lessons Learned by State and Local Governments." *The Public Manager* 23(3):41–44.

Fabry, B. 1998. "Program Evaluation as Quasi-Experimental Research." In *Corporation for Standards and Outcomes.* Vol. III. *Perspectives on Outcomes.*

Frederick, J., and Nicholson, H. 1991. *Assess for Success*. Indianapolis, IN: Girls Incorporated.

Friedman, M. May 1997. *A Guide to Developing and Using Performance Measures in Results-Based Budgeting*. Washington, DC: The Finance Project.

Gambone, M.A. 1997. *Launching a Resident-Driven Initiative. Community Change for Youth Development (CCYD): From Site Selection to Early Implementation*. Philadelphia: Public/Private Ventures.

Gambone, M.A., and Arbreton, J.A. 1997. *Safe Havens: The Contributions of Youth Organizations to Healthy Adolescent Development*. Philadelphia: Public/Private Ventures.

General Accounting Office (GAO). 1997a, May. *Managing for Results: Analytic Challenges in Measuring Performance.* GAO Document No. HEHS/GGD-97-138. Washington, DC: GAO.

GAO. 1997b, May. *Agencies' Strategic Plans Under GPRA: Key Questions to Facilitate Congressional Reviews*. Version 1. GAO Document No. GGD-10.1.16. Washington, DC: GAO.

GAO. 1995, May. *Managing for Results: Steps for Strengthening Federal Management*. GAO Document No. T-GGD/AIMD-95-158. Washington, DC: GAO.

GAO. 1994, December. *Managing for Results: State Experiences Provide Insights for Federal Management Reform.* GAO Document No. GGD-95-22. Washington, DC: GAO.

Government Performance and Results Act (GPRA) of 1993, P.L. 103-62 (107 STAT. 285).

Hatry, H.P. 1997. "Where Rubber Meets the Road: Performance Measurement for State and Local Public Agencies." In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:31–44.

Hatry, H., van Houten, T., Plantz, M.C., and Greenway, M.T. 1996. *Measuring Program Outcomes: A Practical Approach.* Alexandria, VA: United Way of America.

Hwalek, M., and Minnick, M.E. 1997. *Girl Scout Outcomes Study Procedures Manual.* New York: Girl Scouts of the U.S.A.

Joyce, P.G. 1997. "Using Performance Measures for Budgeting: A New Beat, Or Is It the Same Old Tune?" In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:45–62.

Kirk, R.S. 1994. "Some Measurement Issues To Consider When Developing Outcome Measures for Child Welfare Services." Presented at the Second National Roundtable on Outcome Measures in Child Welfare Services, San Antonio, TX, April 7–9.

Leffert, N., Benson, P.L., and Roehlkepartain, J.L. 1997. *Starting Out Right: Developmental Assets for Children*. Minneapolis, MN: Search Institute.

Leffert, N., Saito, R.N., Blyth, D.A., and Kroenke, C.H. 1996. *Making the Case: Measuring the Impact of Youth Development Programs.* Report commissioned by the YMCA of the U.S.A. Minneapolis, MN: Search Institute.

Malone, L., and Markley, H. 1998. "Results of Outcome-Focused Research and Evaluation: A Practice-Based Perspective." *Sixth National Roundtable on Outcome Measures in Child Welfare Services: Summary of Proceedings.* Englewood, CO: American Humane Association, 39–51.

National Academy of Public Administration (NAPA). November 1994. "Toward Useful Performance Measurement. Lessons Learned From Initial Pilot Performance Plans Prepared Under the Government Performance and Results Act." A report from the NAPA Advisory Panel on Improving Government Performance. Washington, DC: NAPA.

National Collaboration for Youth. 1997. *Position Statement on Accountability and Evaluation in Youth Development Organizations.* Washington, DC: National Collaboration for Youth.

Newcomer, K.E. 1997. "Using Performance Measurement To Improve Programs." In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:5–14.

Newcomer, K.E. 1996. "Evaluating Public Programs." In Perry, J.L., ed. *Handbook of Public Administration,* 2nd ed. San Francisco: Jossey-Bass.

Office of National Drug Control Policy (ONDCP). 1998. *Performance Measures of Effectiveness.* Washington, DC: ONDCP.

Pecora, P. 1998. " Future Directions in Outcome Measurement." *Sixth National Roundtable on Outcome Measures in Child Welfare Services: Summary of Proceedings.* Englewood, CO: American Humane Association, 95–99.

Perkins, D.F., Haas, B.E., and Keith, J.G. 1997. "Positive Youth Development With Runaway and Homeless Youth." *New Designs for Youth* 13(4):36–41.

Plantz, M.C., Greenway, M.T., and Hendricks, M. 1997. "Outcome Measurement: Showing Results in the Nonprofit Sector." In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:15–30.

Pratt, C.C., Katzev, A., Ozretich, R., Henderson, T., and McGuigan, W. 1998. *Building Results III: Measuring Outcomes for Oregon's Children, Youth, and Families.* Salem, OR: Oregon Commission on Children and Families.

Rossi, P.H., and Freeman, H.E. 1987. *Evaluation: A Systematic Approach.* Thousand Oaks, CA: Sage Publications.

Scales, P.C., and Leffert, N. 1999. *Developmental Assets: A Synthesis of the Scientific Research on Adolescent Development*. Minneapolis, MN: Search Institute.

Schorr, L. 1995. *The Case for Shifting to Results-Based Accountability*. Washington, DC: Center for the Study of Social Policy.

Wholey, J.S. 1997. "Managing for Results: Performance-Based Management in Theory and Practice." Paper prepared for the International Symposium on Performance-Based Management and Its Training Implications, Caserta, Italy, September 24–26.

Wholey, J.S., and Hatry, H.P. 1992. "The Case of Performance Monitoring." *Public Administration Review* 5(6):604–610.

Wholey, J.S., and Newcomer, K.E. 1997. "Clarifying Goals, Reporting Results." In Newcomer, K.E., ed. *Using Performance Measurement To Improve Public and Nonprofit Programs.* Cited in *New Directions for Evaluation* 75:91–98.